Manoj Kumar Ashok

(312)-284-9898 | mashok@depaul.edu | LinkedIn | GitHub | Portfolio | Chicago, IL

Education

DePaul University Chicago, IL
Master of Data Science, Concentrated on computational methods *Jan 2024 - Nov 2025*
Coursework: Data analysis and regression, **Mining Big data**, Advanced Machine Learning, Fundamentals of Data Science, Advanced Data analysis, Database processing for large scale analytics, **Neural Networks and Deep Learning**, **Natural Language Processing**.

Bharathiar University India
Bachelor's in Computer Applications - AI *July 2020 - Nov 2023*
Coursework: Python programming, Data Structures, Intro to AI and ML, statistics, NLP, Computer vision

Technical Skills

Languages & Developer tools : Python, R,SQL, Git, hadoop, Apache spark
Database Management: MySQL, RDBMS (Oracle), Hive, MongoDB, ETL
Cloud & Pipelines: Azure data factory, Google Dataflow
Visualization: Matplotlib, Seaborn, ggplot2, Tableau
Front end : HTML, CSS, Javascript

Experience

Research Assistant – Machine Learning (GANs) *Jan 2025 – Present*
DePaul University (Prof. David Ramsay) *Chicago, IL*
- Spearheaded research on Conditional GANs, integrating IcGAN and RoCGAN architectures to develop advanced AI-driven image enhancement models.
- Boosted model realism by **30%** (evaluated via FID and IS) and reduced training time by **40%** through hyperparameter tuning and pruning.
- Collaborated across departments to design scalable ML pipelines and validate models for real-time deployment in imaging applications.

Software Developer – Data Engineering Team *Jan 2023 – July 2023*
Zoho Corporation *Chennai, India*
- Developed ETL pipelines using Python and SQL to process CRM/Books datasets, increasing data throughput by **20%**.
- Automated Tableau dashboards using advanced SQL queries, reducing reporting lag by **24%**.
- Built KPI-tracking models to identify and resolve customer support inefficiencies, accelerating resolution speed by **10%**.
- Implemented and monitored Airflow DAGs for reliable data orchestration and reduced batch latency by **15%**.

Projects

Predictive Analysis for Credit Limit | *Python, Scikit-Learn, TensorFlow, SQL, Pandas, NumPy* *July 2024*
- Built regression models using Ridge/Lasso, improving credit limit prediction accuracy by **20%**.
- Applied outlier filtering and normalization techniques, reducing anomalies by **25%**.
- Streamlined cloud-based batch jobs via SQL automation, increasing data throughput by **50%**.

Heart Disease Prediction Using ML | *Logistic Regression, Gradient Boosting, MLP, Random Forest* *Mar 2025*
- Achieved **88.6% accuracy** and **93.9% ROC-AUC** on real-world clinical data using ensemble classifiers.
- Boosted recall by **21%** via stratified sampling and Random Forest feature selection.
- Outlined integration framework for real-time deployment on wearable devices and hospital triage systems.

Bankruptcy Prediction Using Ensemble ML | *XGBoost, LightGBM, Random Forest, SMOTE* *Mar 2025*
- Trained ensemble models on 6.8k-record financial dataset, reaching **98.5% accuracy** and **97.6% F1-score**.
- Utilized SMOTE for minority oversampling, increasing recall by **42%** while maintaining precision.
- Designed full ML pipeline from data prep to tuning, deploying a production-ready ensemble model.